Research Proposal

"Development of an AI-based System for Detection of Lung Infections using Chest X-ray Images"

By

Mr. Rahul Singh G. Bisen

Under the Guidance of

Dr. Archana M. Rajurkar

Professor & Head,

Department of Computer Science & Engineering,

MGM's College of Engineering, Nanded

Submitted to

Swami Ramanand Teerth Marathwada University, Nanded

August, 2023

1. Title of the research proposal

"Development of an AI-based System for Detection of Lung Infections using Chest X-ray images"

2. Introduction

When we breathe, our lungs take oxygen from the air and subsequently deliver it to the bloodstream. Oxygen is required by the cells of our body for proper growth and working. People suffering from lung infections experience breathing difficulties, and a large number of the population across the world are afflicted with various respiratory diseases. Tuberculosis (TB) and Pneumonia are two of the most common lung infections worldwide, and they continue to pose a significant threat to global public health. Many undeveloped and under-developed countries have huge mortality rates due to these lung infections. In 2019, WHO (World Health Organization) reported 1.4 million deaths due to TB, while pneumonia caused approximately 2.6 million deaths globally [1, 2].

Medical imaging is considered as complementary diagnostic tool for lung infections in symptomatic and asymptomatic patients. These patients are referred to radiologists for screening procedures like X-ray, Computed Tomography (CT scan), or Magnetic Resonance Imaging (MRI). From these imaging modalities, X-ray is a widely used diagnostic tool for detecting TB and pneumonia [3, 4]. While various diagnostic methods exist, Chest X-ray (CXR) imaging remains a fundamental tool for the initial assessment of lung infections due to its wide availability, cost-effectiveness, and ability to provide valuable insights into the anatomical and pathological changes in the lungs.

However, the interpretation of CXR images is a complex task that requires a high level of expertise, typically provided by trained radiologists. In the countries like India, where there is shortage of radiologists as compared to overall population of the country, it leads to delay in identification of these illnesses which is crucial for effective treatment and management. Also due to enormous population and heavy load of patients in the hospitals, there are

chances of misdiagnosis and missed diagnosis. So, there is a need for automated systems to assist healthcare providers in saving their valuable time. In light of these challenges, Artificial Intelligence (AI) based techniques offer a promising solution which assist healthcare professionals in the accurate and efficient detection of lung infections from CXR images.

In this research, we propose an AI-based system that will take chest X-ray images of patients as an input, analyze these images by extracting its features to detect the infection in the lung and finally classify it into – Tuberculosis, Pneumonia and Normal. This will aid radiologists in diagnosing these infections and plan the appropriate treatment. This tool will also be helpful for the new doctors working in the remote areas of our country where medical facilities are scarce.

3. Literature survey

Medical imaging and Artificial Intelligence (AI) have emerged as a valuable tool in analyzing and detecting lung infections in the patients. Researchers from different countries are working on developing AI algorithms that can accurately detect and classify TB and pneumonia from X-ray images. Some of them are discussed below –

Yuting Yang et. al. [5] proposed a method for improving pneumonia recognition accuracy using deep learning models by considering and removing the X-ray image background in 2022. They used a dataset of 5840 X-ray images (4265 – Pneumonia & 1575 – Normal). The dataset was augmented using techniques like flipping, scaling, and rotation to prevent overfitting. They developed two deep learning models, ResNet50 and VGG16, to classify X-ray images into two classes: pneumonia and normal. The models were trained and tested with and without the image background. The accuracy of the proposed approach using the VGG16 model was 95.6%.

Cesar Ortiz-Toro et. al. [6] from Spain presented three textural image characterization methods: radiomics, fractal dimension and the recently developed superpixel-based histon, as biomarkers to be used for training Artificial Intelligence models in order to detect pneumonia in chest X-ray imagesin March 2022. Then three machine learning algorithms were used - K-

Nearest Neighbors, Support Vector Machine and Random Forest for classification and validation. To some extent, this study has limitation as the conducted tests did not aim to differentiate between distinct infiltration patterns or various causes of pneumonia.

S. Jaeger et al. [7] used two feature sets for classifying normal and TB patterns using CXRs. The first feature set consists of 6 descriptors, while the second set contains 594 dimensions. The system achieved an AUC of 87% (78.3% accuracy) and 90% (84% accuracy) on MC Dataset and Shenzhen datasets, respectively, using the support vector machine classifier. From the findings, it is observed that this research work emphasizes the potential of machine-learning approaches for improving TB detection accuracy and the importance of using automated methods for analyzing medical images.

L. Hogeweg et al. [8] proposed a Random Forest classifier and used two datasets, namely Find and Treat and TB-NEAT, consisting of 200 images. Unfortunately, it is found that no accuracy of the results was published in their study. The research article by B. Antin et.al., in 2017 [9], focuses on the use of supervised learning approaches for pneumonia identification in CXRs. They used a dataset called ChestX-ray14 [10] and employed the K-nearest neighbor algorithm for classification. The reported accuracy of their model was an AUC of 0.60. However, it is revealed that the data was not divided in a standard way, the AUC was relatively low and there is a need for further improvements in the model.

Wang et al. [11] presented a novel methodology for the detection and localization of pneumonia using CXRs by implementing Co-operative Convolutional Neural Net-works (co-CNN). The U-Net for lung segmentation and co-CNN for *pneumonia* classification and localization was used. The prediction score of the results was 0.2409, and sensitivity, specificity was 0.89 and 0.78 respectively. However, it is revealed that there is room for improvement in terms of sensitivity and specificity. Verma et al. [12] conducted research on tuberculosis, viral pneumonia, bacterial pneumonia, and nor-mal lung nodule classification using a customized CNN. They utilized data augmentation, annealing, and learning rate variation for image processing. The Shenzen and D. Kermany et al. [13] datasets were used by the

researchers. The overall reported accuracy was 99.01%. However, the precision for bacterial pneumonia was found to be low. Additionally, the study lacked experimental results, and further research is needed to validate the findings.

From the study it is revealed that although significant research has been done in designing an automated system for the diagnosis of lung infections, there is a limited work on implementing a single AI-based system which can classify infections into Tuberculosis, Pneumonia and Normal.

4. Statement of Problem

Lung infections, notably tuberculosis (TB) and pneumonia, continue to be significant contributors to global morbidity and mortality. Early and accurate detection of these infections is paramount for effective treatment and improved patient outcomes. However, the accurate interpretation of chest X-ray (CXR) images, a crucial diagnostic tool for lung infections, remains a complex task that heavily relies on the expertise of radiologists. This human-dependent process introduces subjectivity and the potential for diagnostic variability, especially in regions with limited access to specialized medical professionals.

The challenges associated with manual interpretation of CXR images include the risk of misdiagnosis or missed diagnosis due to the intricate nature of pulmonary conditions and the scarcity of experienced radiologists in certain healthcare settings. Moreover, the disparities in diagnostic accuracy across different healthcare facilities further underline the need for a reliable and accessible solution.

Our proposed research aims to address these challenges by working on chest X-ray images and devising an AI-driven methodology for the detection of lung infections and classify the images into Tuberculosis, Pneumonia, and normal cases.

Main objective of this proposal is described by the statement of the problem, expressed as follows:

The objectives of the proposed research work can be divided into:

1. To develop an algorithm for preprocessing and enhancing chest X-ray images.

2. To propose an approach for feature extraction of X-ray images and detection of Tuberculosis from it.

3. To design a methodology for feature extraction of X-ray images and detection of Pneumonia from it.

4. To develop an effective AI-based system for classification of lung infections into - Tuberculosis, Pneumonia and Normal.

5. To evaluate the performance of the AI-based system and validate the results with the ground truth from expert radiologists.

5. Proposed scheme

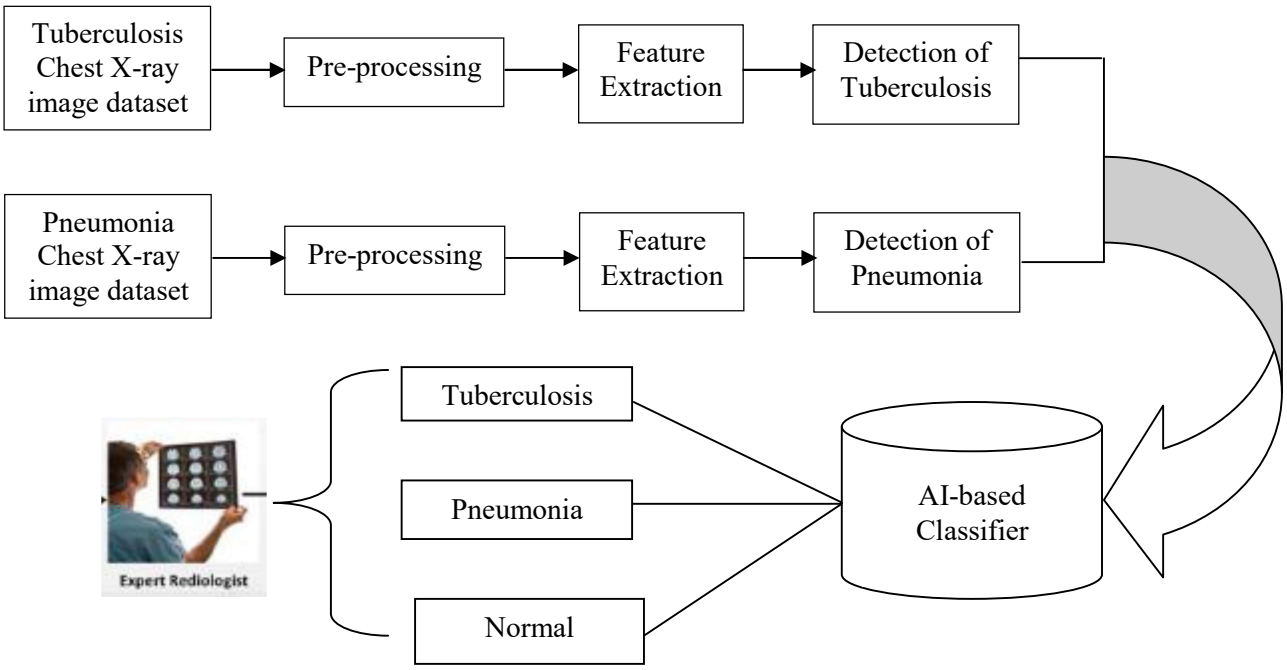

Fig 1: Proposed System Architecture

The proposed system architecture for the research work is shown in the above Fig 1. This work can be divided in the following steps -

1. **Dataset Preparation**: The first step involves collecting and preparing two distinct

datasets. One dataset contains chest X-ray images with cases of Tuberculosis, and the other dataset contains chest X-ray images with cases of Pneumonia. These datasets are crucial for training and testing the AI-based model.

2. **Preprocessing**: The collected chest X-ray images need to undergo preprocessing. Preprocessing steps may include resizing images to a consistent resolution, noise reduction, contrast adjustment, and image enhancement. These steps ensure that the input data is in a suitable format for further analysis.
3. **Feature Extraction**: Feature extraction is an essential step in image analysis. It involves extracting meaningful information or features from the preprocessed images. These features could include textures, patterns, shapes, or other relevant characteristics that can help distinguish between healthy and disease-affected lung tissues. Features can be extracted from various techniques like Gray-Level Co-occurrence Matrix (GLCM), Principal Component Analysis (PCA), Textural Feature Extraction by Radiomics, Fractal Dimension & Superpixel based Histon etc.
4. **Disease Detection**: After feature extraction, the system implements an effective technique or algorithm for the detection of TB and Pneumonia from the X-ray images. In this step the proposed system analyzes the extracted features and detects the presence of these diseases in the chest X-ray images.
5. **AI-Based Classifier**: The AI-based classifier is the core component responsible for making these classifications. The proposed classifier will be trained on a large amount of labeled data, using machine learning or deep learning techniques, to learn the patterns and features associated with Tuberculosis, Pneumonia, and normal lung conditions. The result of this step will classify the images into three distinct classes **-** Tuberculosis, Pneumonia, and Normal.
6. **Validation**: To ensure the reliability and accuracy of the results, the system incorporates a validation step. The results generated by the AI-based classifier will be validated by the expert radiologists. This expert validation helps us to confirm the proposed system's performance and accuracy, ensuring that it is safe and effective for clinical use.

6. Implementation

Python programming or MATLAB with Image Processing Tool box can

be used to implement this proposed work.

A tentative time schedule of the research work planned is given below

Sr. no	Research activity	Schedule
1.	Identifying the problem with more details (Literature Survey)	6 months
2.	Collection and creation of dataset of Chest X-ray images	3 Month
3.	Design of an efficient algorithms for Tuberculosis detection	6 Months
4.	Propose an algorithm for detection of Pneumonia infection	6 Months
5.	Design an algorithm for classifying the images into three classes – Tuberculosis, Pneumonia and Normal	6 Months
6.	Performance evaluation and validate diagnosis results with the ground truth by consulting doctors	3 Months
7.	Thesis write-up, revision and correction	6 Months

7. Conclusion

Tuberculosis and pneumonia are two of the most prevalent and deadly lung infections, with millions of new cases reported each year. According to the World Health Organization (WHO), India is one of the countries with a high burden of Tuberculosis (TB) and Pneumonia. For diagnosing lung infections different kinds of medical image acquisition tools can be used like X-ray, CT scan etc. Out of these, Chest X-ray is economically viable and available in all the remote areas of any country. In addition, the number of doctors is less as compared to the overall population of our country. This leads to delay in diagnosis of these infections. Hence there is a need of an AI-based system which can reduce the burden of doctors and can aid them in diagnosis.

The objective of this proposal is to develop an AI-based system that can accurately detect the presence of tuberculosis and pneumonia in patients by analyzing chest X-ray images. This proposed system will perform preprocessing, feature extraction and detect Tuberculosis and Pneumonia from the dataset. Then an AI-based classifier will classify the images into - Tuberculosis, Pneumonia and Normal. The proposed system aims to assist medical professionals in diagnosis and plan more effective treatment of patients with lung infections. Furthermore, it will also serve as

a valuable resource for the inexperienced doctors in the remote areas of the country where there are limited medical facilities.

References

[1] World Health Organization, Global tuberculosis report 2020, WHO; 2020. https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2020.

[2]World Health Organization, Pneumonia, WHO; 2021.https://www.who.int/news-room/fact-sheets/detail/pneumonia.

[3]B van Ginneken, B M Romeny, M A Viergever, "Computer-aided diagnosis in chest radiography: A Survey". IEEE Trans Med Imaging. 2008;27(4):525-550. doi: 10.1109/TMI.2007.909227

[4] Mandell LA, Wunderink RG, Anzueto A, et al. "Infectious Diseases Society of Ameri-ca/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults", Clin Infect Dis. 2007;44 Suppl 2:S27-S72. doi: 10.1086/511159

[5] Yang, Yuting, Gang Mei, and Francesco Piccialli. "A Deep Learning Approach Considering Image Background for Pneumonia Identification Using Explainable AI (XAI)." IEEE/ACM Transactions on Computational Biology and Bioinformatics (2022).

[6] C´esar Ortiz-Toro, Angel García-Pedrero, Mario Lillo-Saavedra, Consuelo Gonzalo-Martín, "Automatic detection of pneumonia in chest X-ray images using textural features", Elsevier (2022), https://doi.org/10.1016/j.compbiomed.2022.105466

[7] S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. K. Singh, S. Antani, G. Thoma, Y. Wang, P. Lu, and C. J. McDonald, "Automatic Tuberculosis Screening Using Chest Radiographs," in IEEE Transaction on Medical Imaging, Vol. 33, No. 2, p.p. 233245, February 2014.

[8] L. Hogeweg, C. I. Sanchez, P. Maduskar, R. Philipsen, A. Story, R. Dawson, G. Theron, K. Dheda, L. Peters-Bax, and Bram van Ginneken, "Automatic Detection of Tuberculosis in Chest Radiographs Using a Combination of Textural, Focal, and Shape Abnormality Analysis", in IEEE

Transactions on Medical Imaging, Vol. 34, No. 12, December 2015.

[9] B. Antin, K. Joshua, and M. Emil, "Detecting pneumonia in chest Xrays with supervised learning," Stanford Univ., Stanford, CA, USA,Tech. Rep., 2017.

[10]Wang, X., Peng, Y., Lu, L., Lu, Z., & Bagheri, M. (2017). ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2097-2106.

[11]K.Wang, Z. Xiaohong, H. Sheng, and C. Feiyu, “Automatic detection of pneumonia in chest X-ray images using cooperative convolutional neural networks”, in Proc. Chin. Conf. Pattern Recognit. Comput. Vis. (PRCV), 2019, pp. 328-340.

[12]D. Verma, C. Bose, N. Tufchi, K. Pant, V. Tripathi, and A. Thapliyal, "An effcient frame-work for identification of tuberculosis and pneumonia in chest X-ray images using neural network," Procedia Comput. Sci., vol. 171, pp. 217-224, Jan. 2020

[13]Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter,& Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell, 172(5), 1122-1131.

[14]Shilpa Gite, Abhinav Mishra, Ketan Kotecha, Enhanced lung image segmentation using deep learning, Springer (2022), Neural Computing and Applications, https://doi.org/10.1007/s00521-021-06719-8

[15]Huang, Shih-Cheng, Akshay S. Chaudhari, Curtis P. Langlotz, Nigam Shah, Serena Yeung, and Matthew P. Lungren. "Developing medical imaging AI for emerging infectious diseases." nature communications 13, no. 1 (2022): 1-6

Mr. R. G. Bisen
Research Scholar

Dr. A. M. Rajurkar
Research Guide